Ground Photon Extraction From Photon-Counting LiDAR Data Using Adaptive Cloth Simulation With Terrain Index

Guoping Zhang, Shuai Xing, Qing Xu, Pengcheng Li, Dandi Wang, Xinlei Zhang, and Kun Chen

***Abstract*— Photon-counting light detection and ranging (LiDAR) Ice, Cloud, and Land Elevation Satellite-2 (ICESat-2) enables the drafting of global elevation maps. However, vegetation cover, terrain undulation, and residual noise in signal photons substantially reduce the accuracy of ground photon extraction. Existing ground photon extraction algorithms do not consider the factors influencing photon extraction, and the threshold setting lacks a theoretical basis. This study proposed a photon-extraction algorithm with scenario adaptability. First, the cloth simulation (CS) was adapted with a terrain index (TI) to extract ground photons; based on this, the cloth breakage concept was proposed to remove residual noise. We tested the algorithm in Denali National Park and compared its results with those of other extraction algorithms. The results showed that the TI was robust and consistent with the actual terrain; the adaptive CS achieved the best accuracy and precision under different canopy heights and terrains. The mean absolute error (MAE) and root mean square error (RMSE) of extracted photons were 0.95 and 3.41 m, respectively. This study provides a solution to estimate ground elevation using photon-counting LiDAR data.**

***Index Terms*— Cloth simulation (CS), Ice, Cloud, and Land Elevation Satellite-2 (ICESat-2), photon classification, photon-counting light detection and ranging (LiDAR).**

I. Introduction

HIGH-PRECISION ground elevation data are crucial to understand unprecedented changes in the environment. Light detection and ranging (LiDAR) is able to directly obtain the vertical structure of targets; as such, it has been used widely in topographic mapping. Spaceborne laser altimeters are applied for wide-ranging observations and significantly improve the environmental monitoring ability [1].

The Ice, Cloud, and Land Elevation Satellite-2 (ICESat-2) is the new generation of photon-counting LiDAR launched in 2018. The onboard Advanced Topographic Laser Altimeter System (ATLAS) observes the ground with a 10-kHz laser repetition rate; the footprints obtained were approximately 17 m in diameter and sequentially spaced 0.7 m along track [2]. However, this high-sensitivity LiDAR could be easily triggered by solar radiation, atmospheric scattering [3]. The ubiquitous noise photons are intermixed with the signal photons reflected from the ground targets. To extract ground photons, the photon-counting LiDAR data must be denoised and classified. ICESat-2 ATL03 data provides a metric that assists analysts with signal photon identification; however, the accuracy of ground photon extraction from signal photons requires further study.

Researchers have employed several techniques to try to identify ground in the ATL03 signal photon data stream. A direct ground photon extraction algorithm was used to mark the photons with the lowest elevation as ground photons [4] and densify them using a progressive irregular triangular network [5]. Although this algorithm was feasible, it was susceptible to residual noise photons, which are mislabeled as signal photons, but most likely noise photons. To reduce the influence of noise, a moving window and elevation quantile were introduced into the ground photon extraction. Different elevation quantile ranges were used to extract the ground photons according to the difference in land cover and acquired time [6]. This algorithm reduced the influence of noise but lacked terrain considerations. Therefore, a framework containing empirical mode decomposition (EMD) [7] was used to extract the ground photons; however, the extracted results of steep areas were always missing. The classification effect must be improved when dealing with complex terrain data.

As opposed to the algorithms discussed directly above, cloth simulation (CS) [8] is a surface-based classification algorithm in which cloth hardness directly affects the results. The cloth hardness is user-provided, and the original CS cannot satisfy the need for automatic processing of extensive, dense ICESat-2 data. This study utilized CS and adapted it to identify ground photons more accurately in the ATL03 signal photon data stream. CS software control features were developed which automatically, quantitatively describe and adjust for local terrain features. The terrain index (TI) was defined as the ratio of the elevation range of signal photons among different along-track distances and was used to replace cloth hardness.

Manuscript received December 16, 2021; revised February 11, 2022 and March 2, 2022; accepted April 22, 2022. Date of publication April 25, 2022; date of current version May 9, 2022. This work was supported in part by the National Natural Science Foundation of China under Grant 41876105 and Grant 41371436 and in part by the Foundation of Science and Technology on Near-Surface Detection Laboratory under Grant TCGZ2017A008. *(Corresponding author: Shuai Xing.)*

Guoping Zhang is with the Institute of Geospatial Information, PLA Strategic Support Force Information Engineering University, Zhengzhou 450001, China, also with the Collaborative Innovation Center of Geo-Information Technology for Smart Central Plains, Zhengzhou, Henan 450001, China, and also with the Science and Technology on Near-Surface Detection Laboratory, Wuxi 214035, China (e-mail: zaczhang1996@gmail.com).

Shuai Xing is with the Institute of Geospatial Information, PLA Strategic Support Force Information Engineering University, Zhengzhou 450001, China, and also with the Science and Technology on Near-Surface Detection Laboratory, Wuxi 214035, China (e-mail: xing972403@163.com).

Qing Xu, Pengcheng Li, Dandi Wang, Xinlei Zhang, and Kun Chen are with the Institute of Geospatial Information, PLA Strategic Support Force Information Engineering University, Zhenghou 450001, China (e-mail: 13937169139@139.com; lpclqq@163.com; wdd_93@163.com; zxl9602@163.com; ck_u_u@163.com).

Digital Object Identifier 10.1109/LGRS.2022.3170296

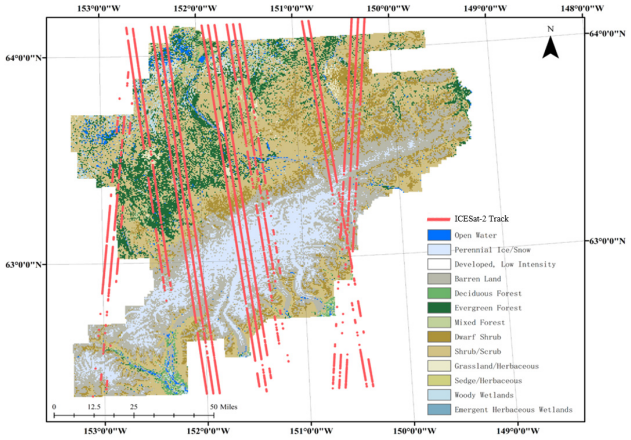

Fig. 1. Study area in the Denali National Park, in inland Alaska, with the track of ICESat-2 data.

In addition, cloth breakage was proposed to avoid the residual noise effect. The performance of the algorithm was evaluated using ICESat-2 data from the Denali National Park. The results show that adaptive CS may accurately and robustly extract ground photons.

II. Study Area and Data

A. Study Area

The study area was Denali National Park (62.29°N–64.07°N, 148.78°W–152.88°W), located in inland Alaska, covering an area of 9492 km^2. The elevation difference of the national park was 6120 m. Cover types vary from spruce bogs to upland spruce-hardwood forests, low shrubs, tundra, and ultimately bare land and snowfields at the highest elevations.

B. ICESat-2 ATL03 Data

ICESat-2 offers 21 data products (ATL00–ATL21). The ATL03 (global geolocated photons) provides the time, longitude, latitude, height, and signal confidence label of each photon, and is the only photon information source required by land-vegetation along-track product and other advanced products [2]. Fig. 1 shows that 16 data pieces were collected in May and June 2019 (https://search.earthdata.nasa.gov/search). Photons with signal confidence labels exceeding two were marked as signals.

C. Reference Data

In Situ data were acquired using Goddard's LiDAR, hyperspectral, and thermal imager (G-LiHT); an airborne measurement system. Using Rigel's VQ-480 LiDAR, G-LiHT generated a 1 m resolution digital terrain model (DTM) and canopy height model (CHM) [9]. As the G-LiHT data acquired in July and August 2014, and July 2018, coincided with ICESat-2 data, 22 datasets were downloaded (https://glihtdata.gsfc.nasa.gov/) and used as a reference. As the study area was inaccessible, the time interval between G-LiHT and ICESat-2 did not create obvious terrain changes.

III. Photon-Extraction Method

When the cloth covers the inverted point cloud, its shape is a DTM. The simulated cloth is composed of a group of particles

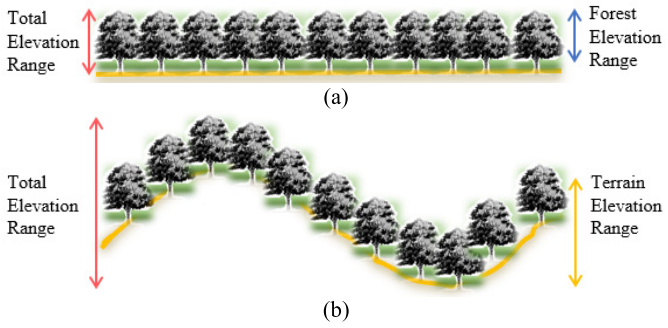

Fig. 2. Typical measurement scenarios of ICESat-2. (a) Flat area. (b) Steep area.

and springs, and the simulation was conducted by iterating the particle positions. In each iteration, the cloth particles first fall under gravity and then bounce back under elasticity.

The rebound distance is determined by the user-provided cloth hardness [8], as shown in (1), which decreases with an increase in terrain undulation

$$\mathrm{RD} = \mathrm{ED}_{\mathrm{particle}} * \left(1 - \frac{1}{2^{\mathrm{hardness}}}\right) \tag{1}$$

where RD is the rebound distance; and $\mathrm{ED}_{\mathrm{particle}}$ is the elevation difference between particles. The greater the hardness, the smaller the rebound of particles and the lower the change in cloth shape. When hardness values were 1, 2, and 3, the moving distances were 1/2, 3/4, and 7/8 of the elevation difference between the cloth particles, respectively.

A. Adaptive CS With TI

The CS is suitable for processing ICESat-2 data, wherein photons are distributed in the plane along the track direction. However, the user-provided cloth hardness in CS cannot meet the demands of vast data automatic processing. As such, the TI was designed to automatically adjust the rebound ratio of cloth particles.

The elevation difference was composed of terrain and ground objects (see Fig. 2), in particular: 1) the elevation difference in flat areas was mainly contributed by ground objects, while the elevation difference in steep areas was mainly caused by terrain and 2) the elevation difference over a short distance was mainly caused by ground objects, whereas the elevation difference over a long distance was mainly caused by terrain.

Thus, comparing the elevation difference at different distances may reflect terrain undulation, although the elevation difference of ground objects is difficult to obtain in advance. Therefore, the ICESat-2 data were divided into the short and long segments, respectively, and the TI [10] was defined as described in the following:

$$\mathrm{TI} = \frac{\mathrm{ED}_l - \mathrm{ED}_s}{\mathrm{ED}_l} \tag{2}$$

where ED_s is the elevation difference of the short segment; and ED_l is the elevation difference of the long segment.

The TI was positively correlated with cloth hardness. With an increase in terrain undulation, the cloth hardness and TI decrease, and the simulated cloth becomes softer.

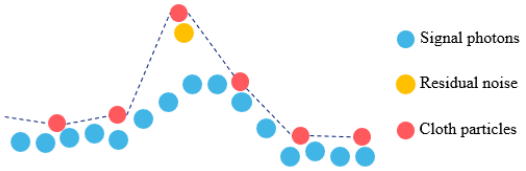

Fig. 3. Illustration of residual noise and cloth damage.

Therefore, TI was used to calculate the rebound distance of cloth particles, as shown in the following:

$$\mathrm{RD} = \mathrm{ED}_{\mathrm{particle}} * \mathrm{TI}. \tag{3}$$

B. Cloth Damage and Residual Noise Detection

When a cloth covers a needle, the gravity of the cloth particles distorts the cloth, rather than describing the exact shape of the needle tip. Similarly, as shown in Fig. 3, when the simulated cloth covers the residual noise (i.e., a noise photon incorrectly identified as signal) far from the local signal photons, it suggests that the cloth is damaged. On this basis, the cloth damage is proposed to detect and remove the residual noise photons.

As residual noise can cause mutations in cloth particle elevation, we calculated the elevation difference between the particles and their average (μ) and variance (σ^2). According to the 3σ principle, when the elevation difference between particles exceeds $\mu+3\sigma$, the cloth will be damaged. Therefore, the photons at the breakage position were marked as noise photons, removed, and the positions of the cloth particles were recalculated.

C. Extraction Method

The major steps of the proposed classification algorithm are as follows.

1) *Signal Photon Segmentation and TI Calculation:* Signal photons were divided into short and long segments and the TI of each short segment was calculated using (2).
2) *Parameter Initialization:* In addition to TI, three auxiliary parameters [8], [10] were also assigned. The first auxiliary parameter, the cloth particle spacing, defined the distance between adjacent cloth particles along track. The second parameter was the moving distance of particles under gravity in each iteration, and the third was the maximum elevation difference between the cloth particle and the corresponding photon.
3) CS and initial ground photons output.
4) Residual noise detection, cloth breakage, and final ground photon output.

IV. Results and Discussion

As different combinations of long and short data segments directly change TI, it is necessary to compare the consistency between TI and the actual terrain under different combinations, investigate the robustness of TI, and identify the best segment combination for the study area. The effectiveness of the adaptive CS was qualitatively and quantitatively verified and compared with classical algorithms. Then, the performances of the adaptive CS under different slopes and vegetation heights were analyzed.

TABLE I
Correlation Between TI and Normalized G-LiHT DTM

Correlation (%)		Short segment (m)					
		50	60	70	80	90	100
Long segment (m)	500	81.91	81.93	80.85	79.22	76.58	75.98
	600	**90.01**	86.29	83.37	87.35	79.93	77.05
	700	87.12	**91.31**	**91.56**	**91.44**	84.47	80.61
	800	81.52	**90.90**	**95.28**	**92.55**	91.61	83.18
	900	74.98	88.35	**91.42**	**92.58**	**90.92**	87.49
	1000	72.13	83.47	87.70	89.91	**90.49**	86.64

A. Verification of TI

To verify the consistency between TI and the actual terrain, the correlation between TI and normalized G-LiHT DTM was determined; based on this, the correlation under different combinations was calculated to identify the best combination of segments. The verified range of the long data segment was 500–1000 m, and the interval was 100 m, while the range of short-distance segment was 50–100 m, and the interval was 10 m; the results are shown in Table I.

Among the results, the correlation of 11 combinations was $>90\%$, and that of 29 combinations was $>80\%$, which shows that TI is effective and robust. When the short segment was 70 m and the long segment was 800 m, the correlation could best (95.28%) depict the terrain in study area and is used in subsequent ground photon extractions.

B. Verification of Accuracy

The photons with a signal confidence label >2 were considered signals and the ground photons from these were extracted with the adaptive CS. The cloth particle spacing was set to 10 m to extract sufficiently dense ground photons, which could be used as ground elevation retrieval data. According to [8], the moving distance of particles under gravity was set to 9.8 m in each iteration, and the maximum elevation difference was set to 0.3 m.

The results and its local amplifications are shown in Fig. 4; the extracted ground photons visually fit the actual terrain. Specifically, by comparing Fig. 4, the vegetation cover and terrain were found to have little influence on the adaptive CS performance, which indicates that CS after adaptation by TI has terrain perception and adaptation ability. In addition, as the cloth particle spacing was set in advance, the extracted ground photons were uniformly distributed, even on steep mountain tops or valleys. Although there were many residual noise photons under the ground surface in the signal photons, the ground photons extracted by adaptive CS do not contain any of them, indicating that cloth breakage is effective.

However, in Fig. 4(b) and (d), at the positions indicated by the arrows, the lack of signal photons in steep areas and under vegetation cover increased the difficulty of ground photon extraction. The correspondence between the cloth particles and the signal photons are set based on the distance between them, cloth particles correspond to the closest signal photons [8].

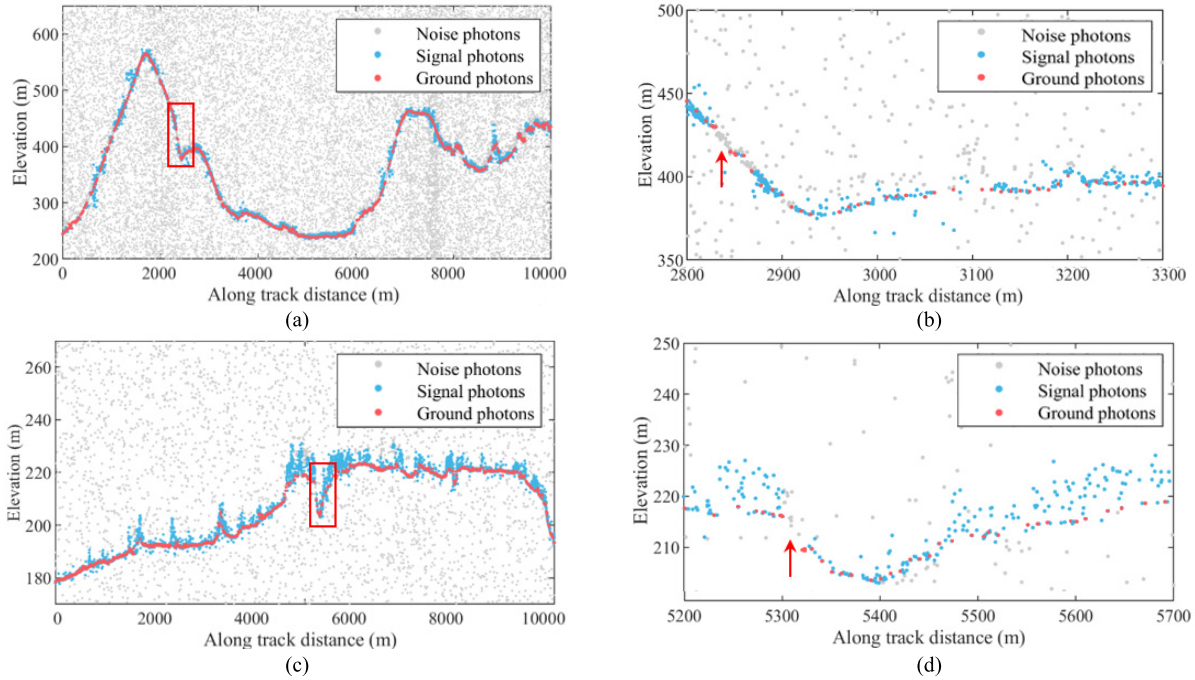

Fig. 4. Results of ground photon extraction and its local amplification (a) bare land in steep terrain and (b) its local amplification at 2800–3300 m; (c) vegetated area in steep terrain and (d) its local amplification at 5200–5700 m.

When signal photons in the vertical direction of the cloth particles are missing, other corresponding signal photons must be found. In Fig. 4(b) and (d), we can see that the extracted photons are closely distributed on both sides of the signal missing positions, which indicates that these photons instead correspond to cloth particles. Although these extracted ground photons are accurate, their distribution is not uniform owing to the lack of signal photons. Additionally, when faced with steeper terrain, the cloth particles may correspond to non-ground photons, such as vegetation photons, thus leading to errors.

To quantitatively evaluate the adaptive CS performance, we compared the elevation of the extracted ground photons with that of the G-LiHT DTM. As shown in (4)–(6), R^2, mean absolute error (MAE), and root mean square error (RMSE) were calculated. We used the algorithms proposed by [4], [6], and [7] as the comparison algorithms and numbered algorithms II–IV:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (e_i - r_i)^2}{\sum_{i=1}^{n} \left(e_i - \sum_{i=1}^{n} r_i\right)^2} \tag{4}$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |e_i - r_i| \tag{5}$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (e_i - r_i)^2} \tag{6}$$

where n is the ground photon number; e_i is the elevation of the ith ground photon; and r_i is the corresponding G-LiHT DTM elevation.

Ground photons were extracted by adaptive CS and comparison algorithms, and the elevation of the extracted photons was compared with reference data; the scatter plots are shown in Fig. 5. As R^2 exceeded 0.9997, there were strong consistencies between the extracted photons and reference ground elevations. In addition, the fitting curves were close to the 1:1 line, indicating that the adaptive CS and comparison algorithms were effective.

Specifically, algorithms II and III exhibited the lowest precision. The results demonstrate that a lack of terrain adaptability and residual noise detection ability will reduce accuracy. The accuracy of algorithm IV was only slightly worse than that of adaptive CS, demonstrating that algorithms have terrain awareness ability. However, the number of ground photons extracted by algorithm IV ($N = 6952$) was significantly less than that of the adaptive CS algorithm ($N = 8946$). A possible explanation is that because of EMD, algorithm IV extracted fewer ground photons in steep areas. It was clear that adaptive CS can extract sufficient ground photons and has better accuracy and precision.

C. Factors Influencing Adaptive CS

Canopy height and terrain undulation were factors that affected ground-photon extraction [5]. Although adaptive CS achieved the best accuracy, we determined its performance under different canopy heights and slopes. Therefore, we calculated the RMSE of the adaptive CS for different scenarios and constructed boxplots.

The RMSE boxplots for different canopy heights are shown in Fig. 6. The canopy heights were grouped as 0–5, 5–10, 10–15, 15–20, and >20 m. The median RMSE increased from 3.24 to 3.62 m with canopy height; the RMSE range became more concentrated with increased canopy height. Specifically, when canopy height >20 m, the RMSE range was <0.8 m. After analysis, adaptive CS was minimally affected by the canopy because the extraction concept of CS is to invert the signal photons and place the simulated cloth on top of the inverted photons. Therefore, the increase in median

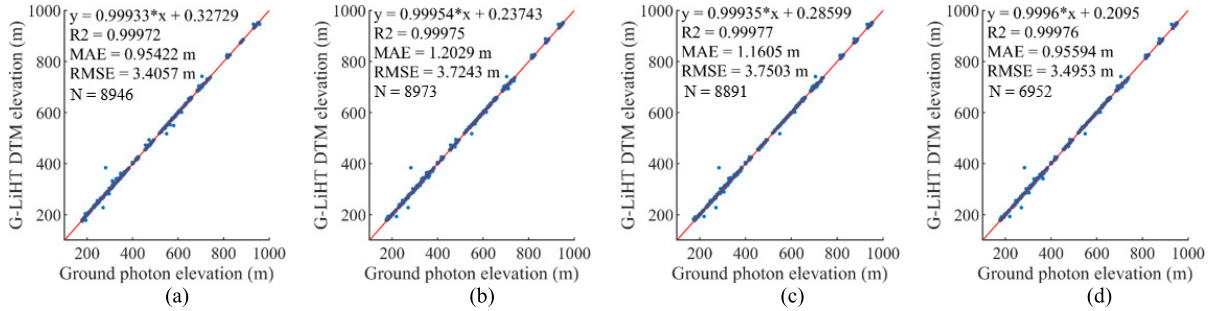

Fig. 5. Scatterplots of ICESat-2-extracted ground photon elevations and G-LiHT DTM elevations from (a) adaptive CS, (b) algorithm II, (c) algorithm III, and (d) algorithm IV.

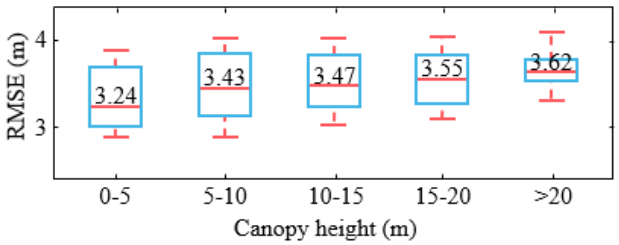

Fig. 6. RMSE boxplot under different canopy height.

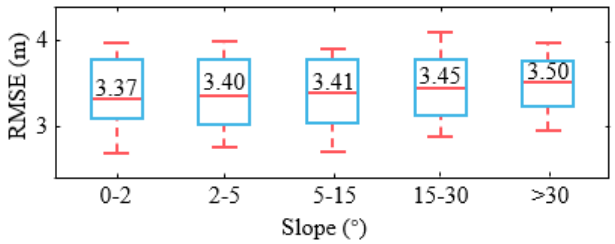

Fig. 7. RMSE boxplot under different slope.

RMSE was more likely to be due to the absence of ground photons under the high canopy [5], and the adaptive CS was robust under different canopy heights.

The RMSE boxplots for different slopes are shown in Fig. 7; slopes were grouped into 0–2, 2–5, 5–15, 15–30, and >30. In Fig. 7, when the slope was 0–2, the precision of the adaptive CS was the highest, and the median RMSE was 3.37 m. When the slope was >30, the median RMSE, at 3.50 m, was greatest. With an increase in the slope, the accuracy of the adaptive CS decreases slightly, which shows that TI is effective, and it allows the simulated cloth to fit the actual ground terrain sufficiently. In addition, with an increase in the slope, the RMSE range also decreased. This demonstrates that cloth breakage may effectively identify and eliminate residual-noise photons in steep areas; the adaptive CS was clearly robust with slope.

V. Conclusion

This study proposed a photon-classification algorithm adapted from CSs. In this algorithm, the automatically set TI was used to adjust the rebound distance of cloth particles as opposed to cloth hardness. Cloth breakage was also proposed to identify and remove residual noise photons.

Tests performed in the Denali National Park show that the accuracy of the adaptive CS was better than that of the existing ground-photon classification algorithms. Furthermore, the adaptive CS was robust to canopy height and terrain undulation. However, the algorithm lacks consideration of the effects of missing signal photons, which may lead to errors in extreme cases. Future research should focus on improving the adaptive CS in this context. Additionally, the performance of the algorithm in more diversified areas should be tested to gain a more comprehensive understanding of CS strengths and limitations.

Acknowledgment

The authors are very grateful for the ICESat-2 data provided by NASA Earthdata and G-LiHT data provided by the Goddard Space Flight Center.

References

[1] D. J. Harding, "ICESat waveform measurements of within-footprint topographic relief and vegetation vertical structure," *Geophys. Res. Lett.*, vol. 32, no. 21, 2005, Art. no. L21S10, doi: 10.1029/2005GL023471.

[2] T. A. Neumann *et al.*, "The ice, cloud, and land elevation Satellite-2 mission: A global geolocated photon product derived from the advanced topographic laser altimeter system," *Remote Sens. Environ.*, vol. 233, Nov. 2019, Art. no. 11325, doi: 10.1016/j.rse.2019.111325.

[3] H. Xie *et al.*, "A comparison and review of surface detection methods using MBL, MABEL, and ICESat-2 photon-counting laser altimetry data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 7604–7623, Jul. 2021, doi: 10.1109/JSTARS.2021.3094195.

[4] M. S. Moussavi, W. Abdalati, T. Scambos, and A. Neuenschwander, "Applicability of an automatic surface detection approach to micro-pulse photon-counting lidar altimetry data: Implications for canopy height retrieval from future ICESat-2 data," *Int. J. Remote Sens.*, vol. 35, no. 13, pp. 5263–5279, Jul. 2014, doi: 10.1080/01431161.2014.939780.

[5] S. Nie *et al.*, "Estimating the vegetation canopy height using micro-pulse photon-counting LiDAR data," *Opt. Exp.*, vol. 26, no. 10, p. A520, Apr. 2018, doi: 10.1364/OE.26.00A520.

[6] S. Popescu *et al.*, "Photon counting LiDAR: An adaptive ground and canopy height retrieval algorithm for ICESat-2 data," *Remote Sens. Environ.*, vol. 208, pp. 154–170, Apr. 2018, doi: 10.1016/j.rse.2018.02.019.

[7] X. Zhu, S. Nie, C. Wang, X. Xi, and Z. Hu, "A ground elevation and vegetation height retrieval algorithm using micro-pulse photon-counting LiDAR data," *Remote Sens.*, vol. 10, no. 12, p. 1962, Dec. 2018, doi: 10.3390/rs10121962.

[8] W. Zhang *et al.*, "An easy-to-use airborne LiDAR data filtering method based on cloth simulation," *Remote Sens.*, vol. 8, no. 6, p. 501, Jun. 2016, doi: 10.3390/rs8060501.

[9] B. Cook *et al.*, "NASA Goddard's LiDAR, hyperspectral and thermal (G-LiHT) airborne imager," *Remote Sens.*, vol. 5, no. 8, pp. 4045–4066, Aug. 2013, doi: 10.3390/rs5084045.

[10] P. Wan *et al.*, "A simple terrain relief index for tuning slope-related parameters of LiDAR ground filtering algorithms," *ISPRS J. Photogramm. Remote Sens.*, vol. 143, pp. 181–190, Sep. 2018, doi: 10.1016/j.isprsjprs.2018.03.020.